Origins and Evolution of Language
Week 2 tutorial briefing
Tutor notes

Comments for tutors are in italics.

Basics of evolutionary theory

The first tutorial will be hands-on, to give you a chance to check your understanding of the basics of the comparative method and evolution by natural selection.

Phylogenetic trees

Using the Evolution Lab game (click "play game"), complete at least the three training trees (Red, green and gecko; Familiar faces; Tree of life: Vegetarian edition). The aim here is to understand why closely related species might be expected to share many traits, and how patterns of shared and differing traits between organisms can be informative about evolutionary history of traits and species (e.g. the patterns of relatedness between species, and when certain traits are likely to have evolved).

I would suggest splitting them into several small groups to work on this exercise and the next one, and either go round and talk to the individual groups as they work through, or reconvene as a group after each task.

They need to use the magnifying glass to examine traits to do this task sensibly.

Things to learn:

- *Species inherit traits from their common ancestor, so if two species share a trait they are likely to share a common ancestor who had that trait. NB They didn't "come from" that common ancestor though - something that people often get confused about is whether humans "come from" monkeys or chimps; no, we share a common ancestor with them, who is likely to have had the traits we share with chimps.*
- *The distribution of traits can therefore help us understand the relationships between species (species that share more traits probably have a more recent common ancestor)* ***and*** *the evolution of those traits (specifically, which traits are recent innovations versus ancestral traits; how recent are those traits).*
- *The tutorial is also designed to help you understand that intuitive groupings (e.g. grouping mushrooms and palm trees together in the first practice tree; grouping dog and fish together in the second; grouping banana and lemon together in the 3rd) may be wrong if they don't capture the pattern of shared traits.*
- *People might ask why you can't have reversals or independent evolution of traits – e.g. in the mushroom-gecko-palm tree tree, couldn't "heterotrophic" evolve twice, once in geckos and once independently in mushrooms. Yes it could, but the idea is that in general the evolutionary tree that involves the fewest independent innovations and independent losses etc is most likely to be right.*
- *One weird thing about the way they have set up these trees is you can't put multiple traits on the root node – e.g. in the same mushroom-gecko-palm tree, I wanted to put "heterotrophic" on the root note and then have it replaced on the palm node. I think*

that’s just a design choice on the app, you can have loss of traits as well as gain of traits in the real world!

- *Hopefully they will make the connection to linguistics and the reconstruction of protolanguages and language phylogenies through shared traits themselves, but if not you can point this out – in fact linguists were the first ones doing comparative reconstruction! See e.g. https://academic.oup.com/sysbio/article/54/4/513/2842862/*

Evolution by natural selection

If the students come out of this exercise understanding that, in general, fitter variants win out, but that this can take a long time or be prone to an element of luck, that is enough.

Use the AlleleA1 web app to answer the questions below. The aim here is to get a basic understanding of the effects of selection and also genetic drift (changes in gene frequency driven by chance).

Drift is the consequence of stochastic events in finite populations where there is a random component to fitness – e.g., even if, on a very long term average, variant A1A1 has higher fitness than other variants and is more likely to survive and reproduce, it can be unlucky and be killed off. You may need to discuss this in relation to Q3-Q5.

The AlleleA1 app allows you to simulate the evolution of a single gene in an imaginary population of organisms. There are two possible genetic variants, called A1 and A2; each organism has two parents and inherits a variant from each, so an individual might be characterised as A1A1 (inherits the A1 variant from both parents), A1A2 (inherits the A1 variant from one parent and the A2 variant from the other), or A2A2 (inherits the A2 variant from both parents). The app allows you to manipulate selection in favour of each possible genotype (A1A1, A1A2, A2A2), by manipulating the relative fitness of each combination of genetic variants.

These models are intentionally abstract – but if you want some examples of easy-to-spot human traits that are controlled by a single gene, see e.g. https://bio.libretexts.org/Bookshelves/Introductory_and_General_Biology/Unfolding_the_Mystery_of_Life_-_Biology_Lab_Manual_for_Non-Science_Majors_(Genovesi_Blinderman_and_Natale)/08%3A_Human_Genetics_and_Cytogenetics/8.02%3A_Human_Traits_Determined_by_Single_Genes. A classic example is eye colour – one gene, two alleles (B and b), each individual is either BB, Bb, bB or bb; individuals with bb have blue eyes, all the other combinations lead to brown eyes (because variant B is dominant).

By default the app starts with all genotypes having equal fitness (they all have relative fitness of 1, so they all produce the same number of offspring), and it plots the frequency of A1 variants in an infinitely large population for 50 generations of evolution; in this scenario, no gene frequencies every change, so the plot is a very boring straight line.

- You can change the course of evolution in this simulated population by changing the relative fitness of the three genotypes, by changing the 3 numbers in the box "Selection — relative fitnesses". Relative fitness is an easy way of modelling differences in the number of offspring each genotype produces on average - a genotype with relative fitness of 1 leaves twice as many offspring as a genotype with relative fitness 0.5.

  *Hopefully it's not too confusing that this is **relative** fitness – it would be easier if they had coded this up so that e.g. you specify how many offspring each variant has. The key thing is higher relative fitness means more offspring.*

  Q1: Can you find relative fitness settings where genetic variant A1 takes over the population (i.e. the line moves to Frequency of Allele A1 close to 1) or dies out (i.e. the line moves to Frequency of Allele A1 close to 0)?

  To make A1 take over, set the relative fitness of organisms with any A2 genes (i.e. A1A2 and A2 A2) to less than 1. E.g. if I leave A1A1 at 1 and set the others to 0.75, allele A1 takes over in 20 generations or so. To make A2 take over, do the reverse.

  Note that the curve is very very smooth, because this is modelling an infinitely large population, where all the maths behaves beautifully and everything behaves exactly according to the relative fitnesses.

  The reset button in the top right is useful by the way!

  Q2: What affects the speed with which this happens? You might want to change to plotting more than 50 generations to see what is happening.

  Basically the bigger the fitness difference the faster it happens. Even a small fitness advantage can eventually pay off – e.g. if I set A1A1 to 1, the others to 0.99 (so a small difference!) then A1 will eventually win out, but it will take 600 generations or so.

  Note also that if you have e.g. A1A1 as fitness 1, A2A2 with fitness 0.5, but leave A1A2 as fitness 1, it takes a long time to kill of allele A2 and get the line all the way to the top -this is because the A2 allele can "hide" in heterozygote A1A2 individuals, selection can't weed them out there. This is like the recessive b variant coding eye colour lurking in the genes of brown-eyed people, to be expressed in their children who happen to inherit the invisible b variant from both parents.

- Things get more interesting if you move from looking at infinitely large populations (where the maths is very neat and the lines very smooth) to populations of a finite size (i.e. where there are a set number of organisms at each generation). You can change this by changing the parameter "Number of finite populations to simulate" to some number other than 0 - e.g. if you set it to 5, it will simulate 5 populations for you, each of 100 individuals (and you control the population size in the next box in the app). The graph will plot the frequency of variant A1 in each population, one line per population; you can click “run again” to rerun the simulations, it will be different every time because these involve a random component for who lives, dies and reproduces – **on average** the relative fitnesses are as you specify them in the boxes, but on any given generation a given organism might get lucky and have more offspring than you expect, or get unlucky and have fewer.

Q3: Reset the app to the default parameters (button in the top right) and then simulate 5 populations where all variants have equal fitness. What happens, and why do you think that happens?

Because we reset the app, this is the situation where all variants have equal fitness, so in the neat infinite population model nothing changes – but here, in finite populations, the lines bounce around as (by chance) A1 individuals have slightly more or slightly fewer offspring than you’d expect, just due to the random element in reproduction. So the populations are not static even in the absence of selection. You could encourage them to zoom out to plot 100 or 200 generations and you’ll start seeing runs where one variant dies out entirely – it goes on an unlucky run and gets down to 0 offspring at some point, at which point it is lost forever. So drift can drive variants to take over or die out, even in the absence of selection.

Q4: What happens if you change the population size to very small (e.g. 10 individuals) or very large (e.g. 1000)?

These drift effects are much more pronounced in small populations, because there is more opportunity for a single unlucky generation to radically change the numbers of a given allele. The bigger the population, the more closely it resembles the behaviour of the smooth, well-behaved infinitely large populations we looked at earlier. Might be worth mentioning that at times in our past the global human population has been ***extremely small****, so just because we are numerous now doesn’t mean we haven’t been prone to genetic drift in the past, and of course if you look at relatively reproductively isolated populations you’d expect lots of drift effects there even in the present day.*

Q5: What if you change the selection parameters as you did before - does the fittest variant always win?

The “run again” button is useful here!

No, the fittest variant doesn’t always win out, because of the stochastic component – so even if I start at equal frequency of variant A1 and A2 and set the relative frequencies of B1 variants a bit lower (e.g. 0.95) then you can still have populations where A1 dies out entirely – the A1 variants just have a run of bad luck. This will happen lots of the time in small populations, because the drift effects are bigger, and only very rarely in large populations. If you make the initial A1 variant initially infrequent (very top box in the interface – set it to e.g. 0.1 or 0.01) this will happen quite often. So evolution by natural selection tends to favour fitter variants, but there *is a stochastic component to it, variants (or species!) can get unlucky and die out even if they should be fitter.*

Optional additional activities

If you get through all of this, they should have read Pinker (2003) for the Monday lecture, so you could ask them what they thought of that!